

- 1 -

TITLE OF THE INVENTION

## DATA SEARCH APPARATUS AND METHOD

5 BACKGROUND OF THE INVENTIONField of the Invention

10 [0001] The present invention relates to a method and apparatus for searching data consisting of binary data and meta-data. The invention also relates to a storage medium in which a control program implementing the above-described method is stored.

Description of the Related Art

15 [0002] Meta-data is "data concerning data", and is also used for describing binary data, such as image data or audio data. By adding meta-data to binary data, image data and audio data is searchable by using keywords. The utility of meta-data has become popular due to the search feature.

20 Accordingly, attempts have been made to use binary meta-data in various formats for conducting the search.

[0003] In a known database, however, if registered binary data to be searched is moved or deleted, such movement or deletion of data is not reflected in the database, thereby failing to obtain a correct search result.

25

SUMMARY OF THE INVENTION

[0004] Accordingly, in view of the above-described problem, it is an object of the present invention to improve ease of operation when conducting a data search by automatically reflecting the addition or deletion of binary data to be searched in the registration status of a database.

[0005] It is another object of the present invention to flexibly respond to internal data by using meta-data described in a data description language, for example, to make it possible to set desired attributes and corresponding values to meta-data by using a data description language even if binary data and meta-data are recorded in the same file, which would otherwise enable the setting of only predetermined attributes and corresponding values to meta-data if the meta-data is described in a binary format.

[0006] In order to achieve the above-described objects, according to one aspect of the present invention, there is provided a data search apparatus for conducting a search of data which includes binary data and meta-data. The data search apparatus includes a database in which the data to be searched is registered. A search unit conducts a search of the data based on a set search condition and the meta-data contained in each item of the data registered in the database. A matching unit matches actually existing data in

the database to the data registered in the database while the search unit is conducting a search.

[0007] According to another aspect of the present invention, there is provided a data search method for conducting a search of data which includes binary data and meta-data. The data search method includes: a searching step of conducting a search of the data based on a set search condition and the meta-data contained in each item of the data registered in a database; and a matching step of matching actually existing data in the database to the data registered in the database during the execution of the searching step.

[0008] According to still another aspect of the present invention, there is provided a storage medium for storing a control program to be executed by a computer. The control program includes: a searching step of conducting a search of the data based on a set search condition and the meta-data contained in each item of the data registered in a database; and a matching step of matching actually existing data in the database to the data registered in the database during the execution of the searching step.

[0009] Further objects, features and advantages of the present invention will become apparent from the following description of the preferred embodiments with reference to the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Fig. 1 is a block diagram illustrating the configuration of an image search apparatus according to a first embodiment of the present invention.

[0011] Figs. 2A and 2B are external views illustrating the image search apparatus shown in Fig. 1.

[0012] Fig. 3 illustrates an example of a graphical user interface (GUI) screen while still image data is being searched.

[0013] Fig. 4 illustrates an example of meta-data to be attached to the still-image data used in the first embodiment of the present invention.

[0014] Fig. 5 illustrates meta-data attached to the still image data used in the first embodiment of the present invention.

[0015] Fig. 6 is a flow chart illustrating a process for searching the still image data by using the GUI screen shown in Fig. 3.

[0016] Fig. 7 is a flow chart illustrating a control process performed by the image search apparatus according to the first embodiment of the present invention.

[0017] Fig. 8 is a flow chart illustrating a process for searching the data which is not registered in a database.

[0018] Fig. 9 illustrates an example of the meta-data shown in Fig. 4 described in the eXtensible Markup Language (XML).

5

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0019] The present invention is described in detail below with reference to the accompanying drawings through illustration of preferred embodiments.

10

First Embodiment

15

[0020] In the following embodiment, meta-data described in a data description language is attached to binary data, and by using such meta-data, the binary data is searched. If there is any binary data which is not registered in a database, such binary data is automatically registered in the database by using the meta-data attached to such binary data. While the search is being conducted, if it is found that binary data registered in the database does not exist in the database, the registration of such binary data is deleted from the directory or the list display portion. With this arrangement, even if the binary data to be searched is moved or deleted, such movement or deletion of data is automatically reflected in the database.

20

25

[0021] Fig. 1 is a block diagram illustrating the

configuration of an image search apparatus according to a first embodiment of the present invention. Referring to Fig. 1, a data input/output unit 100 reads data, such as images, picked up by, for example, a digital camera, from a memory card via a personal computer (PC) card or a universal digital bus (USB). The data input/output unit 100 also writes data, such as images, into a memory card. An input unit 101, which includes a keyboard or a pointing device, inputs data in response to an instruction from a user. The pointing device includes a mouse, a trackball, a tablet, and so on. A storage unit 102 stores binary data, meta-data, etc., and generally, a hard disk is used as the storage unit 102. A display unit 103 displays images, such as a graphical user interface (GUI), and generally, a cathode ray tube (CRT) or a liquid crystal display device is used.

[0022] A central processing unit (CPU) 104 controls the entire processing performed by the above-described elements. A read only memory (ROM) 105 and a random access memory (RAM) 106 provide programs, data, and work areas to the CPU 104 for performing the above-described processing. Control programs used for the processing indicated by the flow charts of Figs. 6 through 8 are stored in the storage unit 102 or the ROM 105. If the control programs are stored in the storage unit 102, they are first read into the RAM 106 and then are executed.

[0023] As for the system configuration, various elements other than the above-described elements are usable. However, since they are not the essential features of the present invention, an explanation thereof is omitted.

5 [0024] Figs. 2A and 2B are external views illustrating the image search apparatus constructed in accordance with the first embodiment of the present invention, and more specifically, Figs. 2A and 2B are a perspective view and a rear view, respectively, of the image search apparatus. A  
10 PC card slot 201 is integrated into the front surface of the image search apparatus, and images are read into the apparatus via a PC card. A power switch 202 is used for turning the image search apparatus on or off. An infrared detector 203 receives a signal from a wireless keyboard 204 or a remote controller 205. The wireless keyboard 204 or  
15 the remote controller 205 forms the input unit 101. On the rear surface of the image search apparatus, a display output terminal 206 is provided, and a display device, such as a CRT or a liquid crystal display device, which forms the  
20 display unit 103, is connected to the display output terminal 206. In addition to the display output terminal 206, a USB terminal for connecting a digital camera to the image search apparatus, or a network connecting terminal for connecting the apparatus to a network may be provided.

25 [0025] In the first embodiment, it is assumed that binary

data to be searched is still image data. Meta-data described in a data description language is attached, as shown in Fig. 5, to all the still image data. In the first embodiment, it is also assumed that the meta-data is described in the eXtensible Markup Language (XML). However, any type of data description language, for example, the HyperText Markup Language (HTML) or the Standard Generalized Markup Language (SGML), is usable.

[0026] In this embodiment, since meta-data is added after the still image data, as shown in Fig. 5, it does not interfere with the processing performed on the still image data (binary data) by a known application. That is, by adding the meta-data after the still image data, the still image data is viewable by using a known application.

[0027] In the first embodiment, the meta-data attached to the still image data is indicated, as shown in Fig. 4, by a pair of data attributes and data values. In the example shown in Fig. 4, as the attributes of the meta-data, five attributes, such as "PhotoGrapher", "Date", "Location", "Event", and "Keyword", are used. As the corresponding data values, the photographer, the time and the date at which the photograph was taken, the place at which the photograph was taken, the name of the event, and the name of the subject, respectively, are indicated.

[0028] Fig. 3 illustrates an example of the GUI screen



while the still image data is being searched. A search condition input unit includes a search attribute input portion 301 and a search keyword input portion 302, and is used for inputting search conditions concerning the meta-data while conducting the search of the still image data. The search attribute input portion 301 is used for specifying the above-described attribute of the meta-data as the search condition, while the search keyword input portion 302 is used for specifying the above-described data value of the meta-data as the search condition. By pressing an inverted delta symbol button 310 located at the right side of the search attribute input portion 301, a list of known attributes is displayable in real time.

[0029] A meta-data display unit includes an attribute display portion 303 and a data value display portion 304. In the attribute display portion 303 and the data value display portion 304, the attribute and the data value, respectively, of a thumbnail image selected from the thumbnails of a list display portion 305 are displayed. By clicking the data value indicated in the data value display portion 304, the corresponding attribute and the data value are designated as the search conditions.

[0030] The list display portion 305 randomly selects n images from the images matching the search conditions, and displays a thumbnail list containing the selected n images.

By clicking a desired thumbnail from the thumbnail list, the corresponding image is selected. In the initial state, since the search conditions are not yet designated, the thumbnails corresponding to the n images selected from all the still image data to be searched are displayed. When the search conditions are specified, the thumbnails corresponding to the n images selected from the still image data matching the search conditions are displayed. In the first embodiment, the maximum number of thumbnails to be displayed in the list display portion 305 at one time is 25. By pressing a reload button 311, subsequent n thumbnails are displayed in the list display portion 305.

[0031] A selected image display portion 306 displays the image selected from the thumbnail images displayed in the list display portion 305. By using a previous button 307 and a next button 308, the thumbnail images disposed before and after the currently selected thumbnail image are selected and viewed. By pressing an OK button 309, the GUI is closed.

[0032] Fig. 6 is a flow chart illustrating a process for searching the still image data by using the GUI shown in Fig. 3. As discussed above, in the initial state, 25 thumbnail images randomly selected from all the still image data to be searched are displayed in the list display portion 305.

[0033] In step S601, a determination is first made as to

whether the user already knows a keyword contained in the meta-data of the target still image data, i.e., whether a keyword to be input is uniquely determined. If the outcome of step S601 is yes, the process proceeds to step S602. In step S602, the attribute is specified by using the search attribute input portion 301, and the search keyword is designated by using the search keyword input portion 302. Accordingly, the specified keyword and the corresponding attribute are set as the search conditions. Then, the process proceeds to step S607 in which the image matching the search conditions is searched.

[0034] If it is found in step S601 that the search keyword is not uniquely determined, the process proceeds to step S603. In step S603, by viewing the thumbnails displayed in the list display portion 305, the user checks for still image data which has been photographed in the same situation as the target still image data. If the result of step S603 is no, the process proceeds to step S604 in which another thumbnail list containing different 25 images is displayed by pressing the reload button 311.

[0035] As described above, after switching the images in the list display portion 305 in step S604, if it is found in step S603 that there is still image data related to the target image data, the process proceeds to step S605. In step S605, by clicking the corresponding thumbnail, the

still image data is selected. Then, the selected still image data is displayed in the selected image display portion 306, and the attributes and the data values of the meta-data of the selected image data are respectively displayed in the attribute display portion 303 and the data value display portion 304. Then, in step S606, by selecting a suitable keyword from the data values displayed in the data value display portion 304, the selected keyword and the corresponding attribute are set as the search conditions. Accordingly, as described above, the search conditions are set without the need to directly input the search keyword. Thereafter, the process proceeds to step S607 in which the image matching the search conditions is searched.

[0036] In step S606, the keyword in the data value display portion 304 is selected by using a pointing device. However, by referring to the keywords displayed in the data value display portion 304, a suitable keyword is directly input into the search keyword input portion 302.

[0037] As discussed above, in step S607, the search is conducted according to the search conditions set in step S602 or S606. Details of the search operation are described below. The 25 thumbnail images randomly selected from the still image data obtained as a search result (matching the search conditions) are displayed in the list display portion 305 as the search result.

[0038] The system process for implementing the above-described processing and operation is shown in, for example, Fig. 7. More specifically, Fig. 7 is a flow chart illustrating the control processing performed by the image search apparatus according to the first embodiment of the present invention. In the following description, data to be searched in the following processing is data consisting of still image data and meta-data, such as that shown in Fig. 5.

[0039] In step S701, the thumbnail images to be searched are displayed in the list display portion 305. As discussed above, 25 still images randomly selected from all the still image data to be searched which are registered in the database are displayed in the list display portion 305. In step S702, it is determined whether the image data to be searched actually exists in the database. If it is found that there is image data which does not actually exist in the database, such data is deleted from the database. Steps S701 and S702 may be executed concurrently, or step S702 may be executed first, and after checking the existence of the thumbnails, the thumbnail images may be displayed. In this case, thumbnail images which do not exist in the database are not displayed. The same applies to steps S705 and S706, and steps S711 and S712.

[0040] In step S703, the process waits for the input of an operation from the operator by using the GUI. If any

operation is performed, the process proceeds to step S704.

[0041] It is determined in step S704 whether the operator has pressed the reload button 311. If the outcome of step S704 is yes, the process proceeds to step S705. In step S705, the data to be searched is selected in a manner similar to step S701, and the corresponding thumbnail images are displayed in the list display portion 305. Then, in step S706, data deletion processing similar to that in step S702 is performed. Then, the process proceeds to step S703 in which the process waits for a new operation from the operator.

[0042] If it is found in step S704 that the operator has not pressed the reload button 311, the process proceeds to step S707. In step S707, it is determined whether the operator has selected a thumbnail image displayed in the list display portion 305, or changed the thumbnail to be displayed in the selected image display portion 306 by pressing the previous button 307 or the next button 308. If the result of step S707 is yes, the process proceeds to step S708. In step S708, the meta-data is extracted from the image data corresponding to the selected thumbnail, and the attributes and the data values of the meta-data are respectively displayed in the attribute display portion 303 and the data value display portion 304. Upon completion of the processing of step S708, the process returns to step

S703, and the process waits for a new operation input from the operator.

[0043] The extraction of the meta-data from the image data in step S708 is performed as follows. In the first embodiment, as shown in Fig. 5, the meta-data described in the XML is attached after the still image data. In this case, the meta-data is described, as shown in Fig. 9.

Accordingly, the portion sandwiched between <Phot> and </Phot> is searched from the back portion of the data file, thereby extracting the meta-data.

[0044] If it is found in step S707 that the thumbnail has not been selected, the process proceeds to step S708 in which it is determined whether the operator has given an instruction to conduct the search by selecting the data in the data value display portion 304 or by inputting the keyword into the search keyword input portion 302. If the outcome of step S709 is yes, the process proceeds to step S710. Although the search attribute input portion 301 and the button 310 are provided for specifying the attribute as the search condition, the attributes of the meta-data are not used for conducting the search in this embodiment.

Accordingly, if it is found in step S709 that the operator has performed any operation on the search attribute input portion 301 or the attribute display portion 303, the process returns to step S703. If the operator has performed

an operation other than those determined in steps S704, S707, and S709, it means that the operator has pressed the OK button 309. Accordingly, the GUI screen is closed. As discussed above, the attributes of the meta-data are not used for conducting the search in this embodiment. This means that the attributes may be safely ignored when conducting the search, though the attribute and the corresponding data value must be input as the search conditions in step S602 for the subsequent steps.

Alternatively, both the attribute and the data value may be used for conducting the search.

**[0045]** In step S710, the database is searched by using the data value of the meta-data specified in step S709 as the search keyword. If there is any data in the files under a predetermined directory which is not registered in the database, a search is also conducted for such data. More specifically, the meta-data contained in the unregistered data under the predetermined directory is checked and compared with the search keyword. Then, if the meta-data matching the search keyword is found, the corresponding image data is also output together with the search result obtained by searching the database. Details of the search operation for unregistered data are described below.

**[0046]** In step S711, according to the search result obtained in step S710, n thumbnail images corresponding to



the found image data are displayed in the list display portion 305. In this case, if there is any thumbnail which does not actually exist, such thumbnail images are deleted from the database in step S712, as in the case of steps S702 and S706.

[0047] If there is any data to be searched which is not registered in the database, such data is registered in step S713. The registration of such data in the database is performed by using the meta-data attached to the still image data. When registering binary data in the database, the file name (path name) of the binary data is registered. As the attribute and the attribute value of the binary data to be registered, the attribute of the corresponding meta-data and the value of the corresponding meta-data (attribute value) are respectively assigned. Upon completion of the processing in step S713, the process returns to step S703 in which the process waits for a new operation input from the operator. Concerning the thumbnails, they are registerable when the corresponding still image data is registered in the database, in which case, the registered thumbnails are displayed in the list display portion 305. If thumbnail images are not registered in the database, they are extracted or generated from the corresponding image data to be displayed.

[0048] Fig. 8 is a flow chart illustrating a process for

the search operation for the data which is not registered in the database. In step S801, the directory in which the data to be searched is stored is checked, and a processing list of the data to be searched is created. Then, in step S802, the counter *i* of the processing list is set to 1. In step S803, the still image data contained in the *i*-th data file of the processing list is expanded. Then, in step S804, the meta-data contained in the expanded data file is extracted. As the XML tags for defining the meta-data, any type of tags may be used, and in this embodiment, the meta-data is partitioned by <Photo> and </Photo>.

[0049] An example of the meta data shown in Fig. 4 described in the XML format is shown in Fig. 9. After extracting the meta-data partitioned between <Photo> and </Photo> in step S804, the process proceeds to step S805. In step S805, the attributes and the data values of the meta-data are extracted and compared with the search conditions. The meta-data may be defined in any format, and in this embodiment, the attribute "str1" and the data value "str2" of the meta-data are described as <ITEM ATTR="str1">str2</ITEM>.

[0050] In step S805, the extracted meta-data is compared with the search condition (search keyword). More specifically, in this embodiment, it is determined whether str2 coincides with the data value of the search keyword.

As discussed above, since only the data value (str2) is compared with the search condition without using the attribute (str1), a more flexible search is conducted. For conducting a more precise search, both the data value (str2) and the attribute (str1) are used and compared with the attribute and the data value of the search keyword, respectively.

[0051] It is then determined in step S806 whether there is any meta-data matched with the search condition. If the outcome of step S806 is yes, the process proceeds to step S807 in which the i-th data is registered in the search result list. Then, it is determined in step S808 whether all the still image data registered in the processing list have been processed. If the result of step S808 is yes, the search processing is completed. If not, the process proceeds to step S809 in which i is incremented by 1, and the process returns to step S803.

[0052] If meta-data is not detected in step S804, it means that meta-data is not attached to the still image data. Accordingly, steps S805 through S807 are omitted, and the process proceeds to step S808. If meta-data is not attached, it is meaningless to register the corresponding still image data in the database. Accordingly, it is preferable that such data (without meta-data) is not registered in step S713.

[0053] According to the above-described processes, the

still image data with the meta-data described in a data description language is searched. If there is any binary data to be searched which is not registered in the database, it is automatically registered by using the corresponding meta-data. While the search is being conducted, if it is found that binary data registered in the database does not actually exist in the database, the registration of such binary data is automatically deleted from the directory or the list display portion. According to this search method, the data registered in the database is moved or deleted while being transparent to the user. This also eliminates the need for the database administrator to register the binary data to be added in the database.

## Second Embodiment

[0054] In the first embodiment, the above-described search method is used for searching the still image data. The present invention is also applicable to moving picture data.

[0055] In the case of moving picture data, the first scene is extracted from the moving picture data, and frames are randomly selected from the first scene. Then, the corresponding thumbnail images are displayed in the list display portion 305. As in the case of the still images, if a moving picture is selected in the list display portion 305,

it is accordingly displayed in the selected image display portion 306. With this arrangement, even if both moving pictures and still images are to be searched, they are handled in a manner similar to the first embodiment.

5 Accordingly, both the still images and the moving pictures are searched without presenting any unnatural feeling to the user.

10 [0056] As the binary data to be searched, instead of the still image data or the moving picture data, audio data may be used. In this case, the selected audio data is played back rather than being displayed in the selected image display portion 306.

15 [0057] In the first and second embodiments, the thumbnails are displayed in the list display portion 305. Alternatively, a file name list is displayed. In particular, if the binary data to be searched is audio data, the corresponding file name list is displayed.

20 [0058] In the foregoing embodiments, a pointing device, such as a mouse, a trackball, or a tablet, is used for specifying the thumbnail in the list display portion 305 or designating the keyword in the data value display portion 304, or for operating the various buttons. Alternatively, a touch panel, which is formed on the display surface of the data search apparatus, or a software keyboard is used.

25 [0059] Instead of using a keyboard, the conditions

(keywords) are directly input into the search attribute input portion 301 or the search keyword input portion 302 by using audio means (which requires speech recognition processing) or barcode means. Alternatively, a software keyboard is used. Alternatively, the keyword is specified with a number assigned to the keyword. In this case, by assigning serial numbers to the individual meta-data when displaying the data values in the data value display portion 304, the keyword is designated with the assigned number.

[0060] The present invention is applicable to a single device (for example, a copying machine or a facsimile) or a system consisting of a plurality of devices (for example, a host computer, an interface, a reader, and a printer).

[0061] The objects of the present invention are also achievable by the following modification. A storage medium for storing a software program code implementing the functions of the first or second embodiment is supplied to a system or an apparatus. Then, a computer (or a CPU or an MPU) of the system or the apparatus reads and executes the program code from the storage medium.

[0062] In this case, the program code itself read from the storage medium implements the functions of the present invention. Accordingly, a storage medium storing such program code constitutes the present invention.

[0063] Examples of the storage medium for storing and

supplying the program code include a floppy disk, a hard disk, an optical disc, a magneto-optical disk, a compact disc read only memory (CD-ROM), a CD-recordable (CD-R), a magnetic tape, a non-volatile memory card, and a ROM.

5 [0064] The functions of the foregoing embodiments are implementable not only by running the read program code on the computer, but also by wholly or partially executing the processing by an operating system (OS) running on the computer or in cooperation with other application software based on the instructions of the program code. The present invention also encompasses such a modification.

10 [0065] The functions of the above-described embodiments are also implementable by the following modification. The program code read from the storage medium is written into a memory provided on a feature expansion board inserted into the computer or a feature expansion unit connected to the computer. Then, a CPU provided for the feature expansion board or the feature expansion unit partially or wholly executes processing based on the instructions of the program code.

15 20 [0066] As described above, according to the present invention, even if binary data to be searched is added or deleted to or from a database, such addition or deletion is automatically reflected in the database, thereby improving the ease of operation for searching the binary data.

[0067] While the present invention has been described with reference to what are presently considered to be the preferred embodiments, it is to be understood that the invention is not limited to the disclosed embodiments. On the contrary, the invention is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims. The scope of the following claims is to be accorded the broadest interpretation so as to encompass all such modifications and equivalent structures and functions.